

# Multimedia Data Mining Approach for Automated Event analysis of suspicious moments

Your Name, Your Co-Author\*

## ABSTRACT

The thirst of knowledge is increasing along with tremendous growth in captured knowledge in form of digital data text, video, image, audio, animation graphics and other sensor data. We can use these multimedia data to get insight into the real world events. In this paper, we are using RFID sensor data as input to statistical model for investigating Evacuation Scenario for abnormal (i.e., explosion) event. Our statistical model assist to get details of location of explosive device, suspicious people, witnesses, victims and details of any causalities.

## Keywords

multimedia data mining, event analysis, surveillance, spatiotemporal analysis,

## 1. INTRODUCTION

Many government organizations, multi-nationals companies, shopping malls, banks and etc are using surveillance system to monitor movements of employees, visitors, machines etc. Ultimate objectives of such surveillance system is to detect suspicious person based on their movements to maintain security and avoid any casualty. Most of these surveillance systems employ video cameras or RFID tags to record locations traversed by each employees and visitors at every time points (i.e., spatio-temporal trajectory). Given such spatio-temporal trajectories, person-in-charge has to quickly discover suspicious movements if any. This requires person-in-charge to continuously monitor many spatio-temporal trajectories at a time, which is certainly not feasible. Furthermore, in case of any abnormal event, many time these spatio-temporal trajectories are useful for further investigation. For example, which movement (i.e., trajectory) was suspicious and why?, etc. Again, it is difficult for human to work with such high voluminous data. This urges a strong

need for automated or semi-automated tool for suspicious behavior detection from trajectories.

There are many different scenarios where we need to investigate abnormal or suspicious behavior. For example, there can be some explosion event followed by evacuation process in office, or some open firing event followed by chaos or some accident or other kind of abnormal events that leads to drastic change in normal behavioral pattern for that environment. In this paper, we are considering the office environment where explosion event is followed by evacuation process. Our goal is to build an intelligent tools that can investigate such evacuation scenarios.

To build such intelligent tool, we need to incorporate similar intuitions the person-in-charge has. In this paper, first we specify set of intuitions human has for the evacuation scenario. Based on these intuitions, we defined two statistical measures and build model to detect any suspicious behavior. Our model is capable to answer when, who, what kind of questions. We validated model on real world RFID based evacuation dataset. These dataset is obtained from ongoing VAST 2008 challenge. Our proposed model is scalable and also flexible. Our contributions are:

- Defined and validated set of intuitions for evacuation scenario
- Propose a scalable and efficient statistical model to detect any suspicious events

In the section 2 we give details of the statistical model derived to analyze the Evacuation scenario. The section 3 will give the experimental details and analysis results. Last section we will conclude with future directions and conclusions.

## 2. INTUITION FOR EVACUATION SCENARIO

To investigate the Evacuation scenario, let us consider the available data and its properties. We are given with the Time( $T$ ) representing intervals between tag readings and each person's location as  $P_{id}(x, y)$  coordinate. Using these spatiotemporal data we can obtain the first order properties like average speed in certain time interval, average traversed area in certain time interval.

Following assumptions and intuitions are use for investigating Evacuation Scenario.

**Intuition 1 [Normal Behavior]:** Ideally there will be normal behavior of people before any abnormal (i.e., explosion) Event in Evacuation Scenario. We can consider normal behavior as,

---

\*Your name is with xyz Department...

- Average speed and Average area traversed by the persons will be low compare to the sudden increase in their speed and traversed area after the abnormal event.
- The average speed and area traversed by persons will remain higher throughout the evacuation process.
- Every one in the area try to reach to their nearest Evacuation places.

**Intuition 2 [Suspicious Behavior]:** Suspicious persons would try to either escape or reach evacuation place before the actual abnormal event happens. We can consider suspicious behavior as,

- Average speed and Average area traversed by the persons will be higher before abnormal event.
- Suspicious person should have visited abnormal event location just prior to the abnormal event occur.
- They might either reach evacuation before others or will escape without entering evacuation area.

**Intuition 3 [Victims Behavior]:** Victims would have normal behavior before abnormal event but will be injured or fainted and ,

- Average speed and Average area traversed by the Victims will be almost near to zero after abnormal event.
- They may not be able to reach to the Evacuation Area.
- They will be found within or very near to abnormal area.

Exploiting above intuitions and given data we can derive a generic statistical model to investigate evacuation scenario.

### 3. STATISTICAL MODEL FOR EVACUATION SCENARIO

In this section we will define the notations and derive the formulations. Following are the basic statistical properties we required to implement evacuation models intuitions.

#### 3.1 Notation and Terminology

Notation	Description
D	Total Duration of taken observations
TNP	Total Number of Persons observed
$P_{id}$	Tag identification of all employees and visitors
$P_{id}(x_t, y_t)$	location coordinate of person $P_{id}$ at time t

Table 1: Notation Table

**Observation Window [W]:** To do the temporal analysis user need to input appropriate observation window parameter for the statistical model.

**Change in location [ $f_{id}(t, t+1)$ ]:** Representing the spatial location of person as (x,y) coordinate in 2D space. To calculate the time window based parameters we need to calculate total number of changed locations within the Time Window. We can define  $f_{id}(t, t+1)$  as follows,

$$f_{id}(t, t+1) = \begin{cases} 1 & \text{if } P_{id}(x_t, y_t) \neq P_{id}(x_{t+1}, y_{t+1}) ; \\ 0 & \text{otherwise.} \end{cases}$$

Where  $P_{id}(x_t, y_t)$  is location of person  $P_{id}$  at time t in 2D space. And  $f_{id}(t, t+1)$  represents change in location as 1 if two consecutive locations are different for person  $P_{id}$  and zero otherwise.

**Average Speed of a Person [ $AS_{P_{id}}$ ]:** For each person  $P_{id}$  we obtain their average speed  $AS_{t_{P_{id}}}$  at time t for time window W as,

$$AS_{t_{P_{id}}} = \sum_t^{t+W} f_{id}(t, t+1). \quad (1)$$

Using above equation we obtain the array  $AS_{P_{id}}$  by accumulating  $AS_{t_{P_{id}}}$  for each time window starting for i=0 to i=D-W, where D is the total time duration of given data.

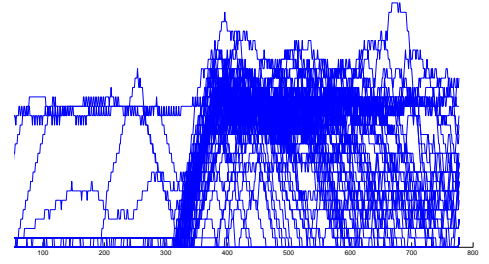


Figure 1: Velocity graph for 60sec time window

**Average Area traversed by a Person [ $AAT_{P_{id}}$ ]:** To obtain average area traversed by a person  $P_{id}$  within time window W at time t is ,

$$AAT_{t_{P_{id}}} = \sum_t^{t+W} U_{id}(t, t+1). \quad (2)$$

Where,

$$U_{id}(t, t+1) = \begin{cases} 1 & \text{if } P_{id}(x_t, y_t) \neq P_{id}(x_{t+1}, y_{t+1}) \\ & \text{and } P_{id}(x_t, y_t) \text{ is unique in W;} \\ 0 & \text{otherwise.} \end{cases}$$

We obtain the array  $AAT_{P_{id}}$  by accumulating  $AAT_{t_{P_{id}}}$  for each time window starting for i=0 to i=D-W.

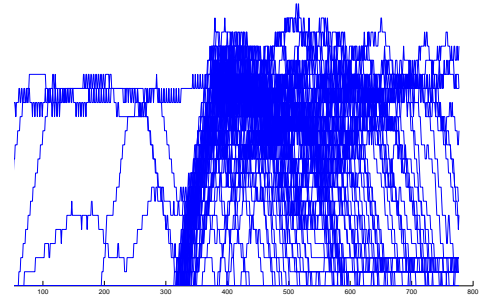


Figure 2: Region Coverage for 60sec time window

**Global Average Speed [ $GAS_t$ ]:** For total number of persons (TNP), the global average speed  $GAS_t$  at time t

for time window  $W$  is,

$$GAS_t = \frac{\sum_t^{t+W} AS_{tP_{id}}}{TNP}. \quad (3)$$

We obtain the GAS by accumulating  $GAS_t$  for each time window starting for  $i=0$  to  $i=D-W$ , where  $D$  is the total time duration of given data.

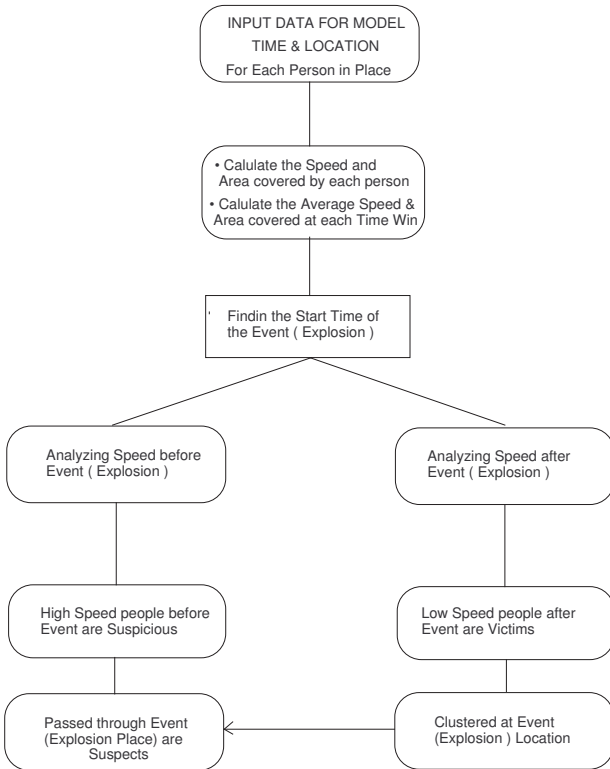
**Global Average Area traversed[GAAT<sub>t</sub>]:** The global average area traversed for total number of persons (TNP) within the time window  $W$  at time  $t$  is ,

$$GAAT_t = \frac{\sum_t^{t+W} AAT_{tP_{id}}}{TNP}. \quad (4)$$

The GAAT obtained by accumulating  $GAAT_t$  for each time window starting for  $i=0$  to  $i=D-W$ . Based on the basic statistical parameters obtained above we will derive mathematical model for Evacuation process analysis.

### 3.2 Model

The Figure3 represents the statistical model for evacuation scenario. It describes the complete flow of computations starting from processing raw data to find the required answers for evacuation scenario. For the given two dimensional spatiotemporal input data we do transformation to obtain the one dimensional representation of data. These data is used for further statistical computation based on intuition to derive the different values representing the specific details of evacuation process. In the following section we will



**Figure 3: Model**

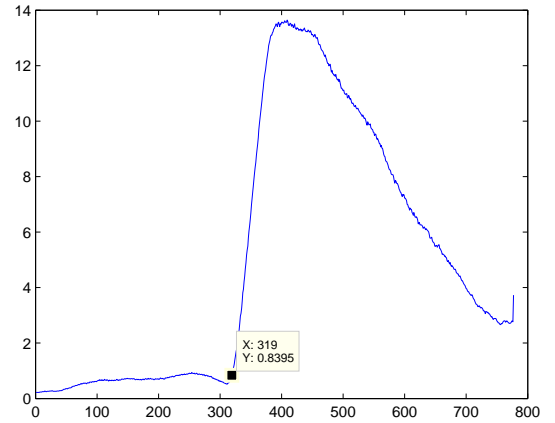
see how we obtain the answers to the different questions for evacuation scenario using the proposed model.

## 4. METHODOLOGY

### 4.1 When?

We derive the starting time of the abnormal event using our basic intuition[1]. We are expecting the drastic change in the global behavior upon occurrence of abnormal event. These behavior can be observed as change in the speed of each person and thus change in global average speed. To use the statistical properties to derive at such starting time we do the following,

- Obtain the Global Average Speed GAS.
- Find the Global Maximum[GM] value from GAS.
- This Global Maximum[GM] value approximately indicates the starting time of abnormal event.



**Figure 4: Global Average Speed per time window over entire time duration**

We can see figure4 of Global Average Speed obtained for real data set. Which justify our intuition of having possible Global Maximum value as indicator of starting time of event.

### 4.2 Where?

Where was the device set off ? is the important question to for investigation. We have categorized the people based on their behavior by intuitions mentioned earlier in the paper. There is strong relation between the victims behavior and the location of abnormal event. Thus, we do following steps to find the location of the explosive device.

- Calculate the Average speed of person  $AS_{P_{id}}$  for duration after abnormal event start.
- Find the distance based outlier for  $AS_{P_{id}}$  comparing with GAS over the time duration after the abnormal event started.
- Found set of  $P_{ids}$  is considered as set of victims.
- Cluster the spatial trajectory  $P_{id}(x, y)$  for  $P_{ids}$  in victims set for duration after abnormal event.
- Centroid of such cluster is potential location for explosive device.

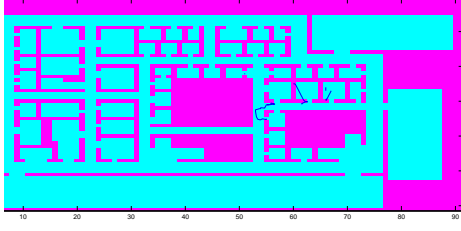


Figure 5: Expected Area of Explosion

The figure 6 shows the potential location for explosive device using the above steps.

Another important information regarding the location is, where is the evacuation area? Which is easily obtained as cluster of spatial trajectory  $P_{id}(x, y)$  of TNP over duration near to end of evacuation. The figure?? shows the evacuation location.

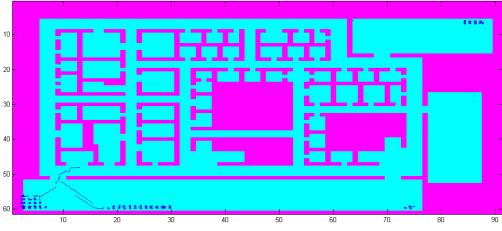


Figure 6: Evacuation Area

### 4.3 Who?

We need to identify set of potential suspects, set of witnesses and set of victims for evacuation scenario. We already found the set of victims during the process of finding device location. Now we derive set of suspects as follows,

- Calculate the Average speed of person  $AS_{P_{id}}$  for duration before abnormal event start.
- Find the distance based outlier for  $AS_{P_{id}}$  comparing with GAS over the time duration before the abnormal event started.
- Found set of  $P_{ids}$  is considered as preliminary set of suspects.
- Using preliminary set of suspects, find the set of  $P_{ids}$  having existence of spatial trajectory  $P_{id}(x, y)$  in abnormal event area for the duration after abnormal event.
- The obtained set of  $P_{ids}$  is potential set of suspects.

The figure?? shows the trajectory of most potential suspect obtained.

Similarly we can find witnesses,

- Considering the set of  $P_{ids}$  which neither belongs to victims set nor to preliminary suspects set.
- Using this set of  $P_{ids}$ , find the set of  $P_{ids}$  having existence of spatial trajectory  $P_{id}(x, y)$  within certain range of abnormal event area for the duration after abnormal event.

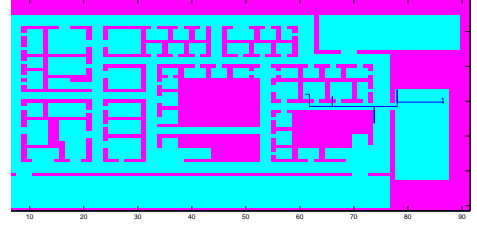


Figure 7: Most suspicious Person's Traversal Path

- The obtained set of  $P_{ids}$  is potential set of witnesses.

### 4.4 To identify any suspects and/or witnesses who managed to escape the building.

We have already identified the set of suspicious people and the witnesses. Now we need to find out the Evacuation Location so that we can find the set of suspects or witnesses who escaped instead of entering the evacuation location. To find the Evacuation place we can cluster the location of people during the end time of event ( explosion). The clusters with threshold density are considered the valid evacuation locations. While the outliers are the potential escape locations.

### 4.5 To identify any casualties

We already obtained the set of distance outliers with very low speed or no speed after the event start time. These are the people who were fainted or dead or injured due to explosion.

## 5. CONCLUSION AND FUTUREWORK

Plotting the Traversal path of all persons for a given input data can be seen in figure8 . The complete path shows clearly the evacuation area. Even the path chosen by people to reach to evacuation area is also clearly visible as thick blue regions. Though this traversal path is not easy to analyze for investigating scenario details like where, when, who etc. Thus, We can see the need for obtained statistical tool to investigate scenarios efficiently.

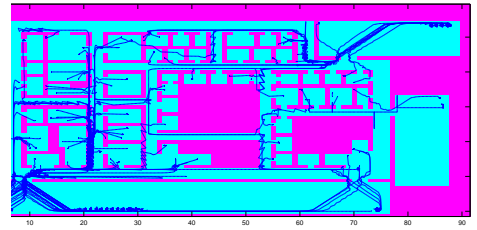


Figure 8: Complete Traversal Path

In the paper we come up with model for spatiotemporal data obtained through RFID surveillance systems. In future we will incorporate other modalities and derive the model for more complex scenarios.